WHAT IS CLAIMED IS:

1. A method for recovering 3D scene structure and camera motion from image data obtained from a multi-image sequence, wherein a reference image of the sequence is taken by a camera at a reference perspective and one or more successive images of the sequence are taken at one or more successive different perspectives by translating and/or rotating the camera, the method comprising the steps of:

(a) determining image data shifts for each successive image with respect to the reference image; the shifts being derived from the camera translation and/or rotation from the reference perspective to the successive different perspectives;

(b) constructing a shift data matrix that incorporates the image data shifts for each image;

(c) calculating two rank-3 factor matrices from the shift data matrix using SVD, one rank-3 factor matrix corresponding to the 3D structure and the other rank-3 factor matrix corresponding tothe camera motion;

(d) recovering the 3D structure from the 3D structure matrix by solving a linear equation; and

(e) recovering the camera motion from the camera motion matrix using the recovered 3D structure.

2. The method of claim 1, wherein the image data is one or more selected from the

group consisting of points, lines and intensities.

3. The method of claim 1, wherein the step of determining image data shifts includes initially recovering and compensating for camera rotation.

4. The method of claim 1, wherein step (b) comprises:

computing H and $\overline{\mathrm{D}}_{CH}$, where H is a $(N_{tot}-3)\times N_{tot}$ matrix and $\overline{\mathrm{D}}_{CH} \equiv \mathrm{C}^{-1/2}\,\overline{\mathrm{D}}\,\mathrm{H}^{\mathrm{T}}$ and $N_{\mathrm{tot}} \equiv 2N_P + 2\,N_L + N_X$, where N_P, N_L, and N_X equal the number of points, lines and intensities, respectively, and C is a $(\mathrm{N_I}-1)$ x $(\mathrm{N_I}-1)$ matrix with $C_{ii'} \equiv \delta_{ii'} + 1$ where $C_{ii'}$ is a constant table of values for all images and $\delta_{ii'}$ is an operator equal to 1 if $i = i'$ and 0 if $i \neq i'$, and $\overline{\mathrm{D}} \equiv [\mathrm{S}\,\omega_{\mathrm{L}}\Lambda\,\omega_{\mathrm{I}}\Delta]$, H^{T} is the transpose of the H matrix, where H is a matrix defined such that H^{T} is an identity matrix and annihilates $\overline{\Psi}_x, \overline{\Psi}_y, \overline{\Psi}_z$ where $\overline{\Psi}_x^T \equiv \left[\Psi_x^T\,\omega_{\mathrm{L}}\Psi_{Lx}^T\,\omega_{\mathrm{I}}\Psi_{Ix}^T\right]$ and similarly for the y and z components where, ω_{I} and ω_{L} are constant weights that the user sets, and where $\Psi_{Lx} \equiv \begin{bmatrix}\{\mathrm{P}_U\cdot(\hat{\mathrm{x}}\times\mathrm{A})\}\\ \{\mathrm{P}_L\cdot(\hat{\mathrm{x}}\times\mathrm{A})\}\end{bmatrix}, \Psi_{Ly} \equiv \begin{bmatrix}\{\mathrm{P}_U\cdot(\hat{\mathrm{y}}\times\mathrm{A})\}\\ \{\mathrm{P}_L\cdot(\hat{\mathrm{y}}\times\mathrm{A})\}\end{bmatrix}, \Psi_{Lz} \equiv \begin{bmatrix}\{\mathrm{P}_U\cdot(\hat{\mathrm{z}}\times\mathrm{A})\}\\ \{\mathrm{P}_L\cdot(\hat{\mathrm{z}}\times\mathrm{A})\}\end{bmatrix}$ where A_l is the unit normal to the plane containing the line and the camera center at the reference image and where $\hat{\mathrm{x}}, \hat{\mathrm{y}}, \hat{\mathrm{z}}$ are unit vectors in the x, y and z directions, and where $\Psi_x \equiv \begin{bmatrix}\{r_x^{(1)}(q)\}\\ \{r_y^{(1)}(q)\}\end{bmatrix}$, $\Psi_y \equiv \begin{bmatrix}\{r_x^{(2)}(q)\}\\ \{r_y^{(2)}(q)\}\end{bmatrix}$, $\Psi_z \equiv \begin{bmatrix}\{r_x^{(3)}(q)\}\\ \{r_y^{(3)}(q)\}\end{bmatrix}$ where $q=(x,y)$ is the image position of the tracked point in the reference image and

the three point rotational flows $r^{(1)}(x,y), r^{(2)}(x,y), r^{(3)}(x,y)$ are defined by $\left[r^{(1)}, r^{(1)}, r^{(1)}\right] \equiv \left[\begin{pmatrix}-xy\\ -(1+y^2)\end{pmatrix}, \begin{pmatrix}1+x^2\\ xy\end{pmatrix}, \begin{pmatrix}-y\\ x\end{pmatrix}\right]$ and where $\Psi_{\mathrm{I}x} \equiv -\{\nabla I\cdot \mathrm{r}^{(1)}(\mathrm{p})\}, \Psi_{\mathrm{I}y} \equiv -\{\nabla I\cdot \mathrm{r}^{(2)}(\mathrm{p})\}, \Psi_{\mathrm{I}z} \equiv -\{\nabla I\cdot \mathrm{r}^{(3)}(\mathrm{p})\}$, and where p gives the image coordinates of the pixel positions, and

where Λ is an $N_I \times 2N_L$ matrix where each row corresponds to a different image i and equals $\left[\{\mathrm{P}_U\cdot\delta\mathrm{A}^i\}^T\{\mathrm{P}_L\cdot\delta\mathrm{A}^i\}^T\right.$ where P_U and P_L are unit 3-vectors projecting onto two directions $\mathrm{A}_l\times(\hat{z}\times\mathrm{A}_l)$ and $\hat{z}\times\mathrm{A}_l$ which we refer to respectively as the upper and lower

directions, where δA^i is the line flow $\delta A_l^i \equiv A_l^i - A_l^0$, and $\hat{z}$ is the unit vector in the z direction and where S is a $N_i \times 2N_p$ matrix where each row corresponds to a different image i and equals $\left[\{s_x^i\}^T \{s_y^i\}^T\right]$, where $s_m^i \equiv q_m^i - q_m^0$ and denotes the image displacement for the m-th tracked point and where Δ is a $N_I \times N_X$ matrix, where each row corresponds to an image i and equals $\{\Delta I^i\}^T$ where ΔI is the change in image intensity with respect to the reference image and where I^i denotes the i-th intensity image and where $I_n^i = I^i(p_n)$ denotes the image intensity at the n-th pixel position in I^i and where the notation $\{V\}$ is used to denote a vector with elements given by the V^a .

5. The method of claim 1, wherein step (c) comprises:

computing the best rank-3 factorization of $\overline{D}_{CH} \approx M^{(3)}S^{(3)T}$ where $M^{(3)}, S^{(3)}$ are rank 3 matrices corresponding respectively to motion and structure, using SVD.

6. The method of claim 1, wherein step (d) comprises:

eliminating structure unknowns Q_z, B_z and Z^{-1} from the $\overline{\Phi}_a$ to get $3N_{tot}$ linear constraints on the U and Ω using the linear equation, $[\overline{\Phi}_x\ \overline{\Phi}_y\ \overline{\Phi}_z] = H^T S^{(3)} U + [\overline{\Psi}_x\ \overline{\Psi}_y\ \overline{\Psi}_z]\Omega$, where U and Ω are unknown 3x3 matrices, and $\overline{\Phi}$ and $\overline{\Psi}$ represent total translational and rotational flow vectors respectively, and solving these constraints with $O(N_{tot})$ computations using the SVD, where the total translational flow vectors are defined by $\overline{\Phi}_x \equiv \begin{bmatrix} \Phi_x \\ \omega_L \Phi_{Lx} \\ \omega_I \Phi_{Ix} \end{bmatrix}$ and similarly for the y and z components, and

$$\Phi_{Lx} \equiv \begin{bmatrix} \{A_x B_U\} \\ \{A_x B_L\} \end{bmatrix}, \Phi_{Ly} \equiv \begin{bmatrix} \{A_y B_U\} \\ \{A_y B_L\} \end{bmatrix}, \Phi_{Lz} \equiv \begin{bmatrix} \{A_z B_U\} \\ \{A_z B_L\} \end{bmatrix}$$

where A is the normal to a first plane that passes through the center of projection and the imaged line in the reference image, and the normal to a second plane, B is defined by requiring $B \cdot A = 0$ and $B \cdot Q = -1$ for any point Q the 3D line L, and where $B_U \equiv B \cdot P_U$ and $B_L \equiv B \cdot P_L$ are the upper and lower components of B, and where

$$\Phi_x \equiv -\begin{bmatrix}\{Q_z^{-1}\}\\ \{0\}\end{bmatrix}, \Phi_y \equiv -\begin{bmatrix}\{0\}\\ \{Q_z^{-1}\}\end{bmatrix}, \Phi_z \equiv -\begin{bmatrix}\{q_x Q_z^{-1}\}\\ \{q_y Q_z^{-1}\}\end{bmatrix}$$

and where

$\Phi_{Ix} \equiv -\{Z^{-1} I_x\}, \Phi_{Iy} \equiv -\{Z^{-1} I_y\}, \Phi_{Iz} \equiv \{Z^{-1}(\nabla I \cdot \mathrm{p})\}$, and

where Q is the 3d coordinate for a 3D tracked point corresponding to an image pixel in the reference image, and Z_n is the depth from the camera to the 3D point imaged at the n-th pixel along the cameras optical axis; and

recovering the structure unknowns Q_z, B_z and Z^{-1} from $\left[\overline{\Phi}_x\ \overline{\Phi}_x\ \overline{\Phi}_x\right] = \mathrm{H}^T S^{(3)} U + \left[\overline{\Psi}_x\ \overline{\Psi}_y\ \overline{\Psi}_z\right]\Omega$, given U and Ω.

7. The method of claim 1, wherein step (e) comprises:

using $S^{(3)}U \approx \left[\overline{\Phi}_x\ \overline{\Phi}_y\ \overline{\Phi}_z\right]$ and

$\overline{\mathrm{D}}_{CH} \approx C^{-1/2}\{T_x\}\overline{\Phi}^T \mathrm{H}^{\mathrm{T}} + C^{-1/2}\{T_y\}\overline{\Phi}_y^T \mathrm{H}^{\mathrm{T}} + C^{-1/2}\{T_z\}\overline{\Phi}_z^T \mathrm{H}^{\mathrm{T}}$ to recover the translations, and

recovering the rotations $\omega_x^i, \omega_y^i, \omega_z^i$ from

$\omega_x^i \overline{\Psi}_{xn} + \omega_{yx}^i \overline{\Psi}_{yn} + \omega_{zx}^i \overline{\Psi}_{zn} = C^{-1/2}\overline{\mathrm{D}}_n^i - \left(C^{-1/2}\left(\{T_x\}\overline{\Phi}^T + \{T_y\}\overline{\Phi}_y^T + \{T_z\}\overline{\Phi}_z^T\right)\right)_n^i$, wherein C is a constant $(\mathrm{N_I}-1)$ x $(\mathrm{N_I}-1)$ matrix with $C_{ii'} \equiv \delta_{ii'} + 1$ and T represents the translation.

8. A method for recovering 3D scene structure and camera motion from image data obtained from a multi-image sequence, wherein a reference image of the sequence is taken by

a camera at a reference perspective and one or more successive images of the sequence are taken at one or more successive different perspectives by translating and/or rotating the camera, the method comprising the steps of:

(a) initially recovering and compensating for camera rotation after warping the reference image by the previously estimated translational displacements;

(a) determining image data shifts for each successive image with respect to the reference image; the shifts being derived from the camera translation and/or rotation from the reference perspective to the successive different perspectives;

(b) constructing a shift data matrix that incorporates the image data shifts for each image;

(c) modifying the image shifts by multiplying the image shifts $\delta A_l^i, s_m^i, \Delta I_n^i$ and redefining these quantities to include the denominator factors $1-Q_z^{-1}T_Z$ for points, $1-Z^{-1}T_z$ for intensities and $1-B\cdot T$ for lines where $s_m^i \equiv \mathrm{q}_m^i - \mathrm{q}_m^0$ and denotes the image displacement for the m-th tracked point, where ΔI_n^i is the change in image intensity and where I^i denotes the i-th intensity image and where $I_n^i = I^i(\mathrm{p}_n)$ denotes the image intensity at the n-th pixel position in I^i and where δA_l^i represents the line flow

(d) calculating two rank-3 factor matrices from the modified shift data matrix using SVD, one rank-3 factor matrix; corresponding to the 3D structure and the other rank-3 factor matrix corresponding to the camera motion;

(e) recovering the 3D structure from the 3D structure matrix by solving a linear

equation; and

(f) recovering the camera motion from the camera motion matrix using the recovered 3D structure.

9. The method of claim 8, wherein the image data is one or more selected from the group consisting of points, lines and intensities.

A3 cont

10. The method of claim 8, wherein step (b) comprises:

computing H and $\overline{\mathrm{D}}_{CH}$, where H is a $(N_{tot}-3)\times N_{tot}$ matrix and $\overline{\mathrm{D}}_{CH} \equiv \mathrm{C}^{-1/2}\ \overline{\mathrm{D}}\ \mathrm{H}^{\mathrm{T}}$ and $N_{\mathrm{tot}} \equiv 2N_p + 2\,N_L + N_X$, where N_p, N_L, and N_X equals the number of points, lines and intensities, respectively, and C is a $(\mathrm{N_I}-1)$ x $(\mathrm{N_I}-1)$ matrix with $C_{ii'} \equiv \delta_{ii'}+1$ where $C_{ii'}$ is a constant table of values for all images and $\delta_{ii'}$ is an operator equal to 1 if $i=i'$ and 0 if $i \neq i'$ and $\overline{\mathrm{D}} \equiv [\mathrm{S}\,\omega_{\mathrm{L}}\Lambda\,\omega_{\mathrm{I}}\Delta]$, H^{T} is the transpose of the H matrix, where H is a matrix defined such that H^{T} is an identity matrix and annihilates $\overline{\Psi}_x, \overline{\Psi}_y, \overline{\Psi}_z$ where $\overline{\Psi}_x^T \equiv [\Psi_x^T\ \omega_{\mathrm{L}}\Psi_{Lx}^T\ \omega_{\mathrm{I}}\Psi_{Ix}^T]$ and similarly for the y and z components and where ω_{I} and ω_{L} are constant weights that the user sets, and where $\Psi_{Lx} \equiv \begin{bmatrix} \{\mathrm{P}_U \cdot (\hat{\mathrm{x}}\times \mathrm{A})\} \\ \{\mathrm{P}_L \cdot (\hat{\mathrm{x}}\times \mathrm{A})\} \end{bmatrix}, \Psi_{Ly} \equiv \begin{bmatrix} \{\mathrm{P}_U \cdot (\hat{\mathrm{y}}\times \mathrm{A})\} \\ \{\mathrm{P}_L \cdot (\hat{\mathrm{y}}\times \mathrm{A})\} \end{bmatrix}, \Psi_{Lz} \equiv \begin{bmatrix} \{\mathrm{P}_U \cdot (\hat{\mathrm{z}}\times \mathrm{A})\} \\ \{\mathrm{P}_L \cdot (\hat{\mathrm{z}}\times \mathrm{A})\} \end{bmatrix}$ where A_l is the unit normal to the plane containing the line L and the camera center at the reference image, and where $\hat{\mathrm{x}}, \hat{\mathrm{y}}, \hat{\mathrm{z}}$ are unit vectors in the x, y and z directions, and where $\Psi_x \equiv \begin{bmatrix} \{r_x^{(1)}(q)\} \\ \{r_y^{(1)}(q)\} \end{bmatrix}, \Psi_y \equiv \begin{bmatrix} \{r_x^{(2)}(q)\} \\ \{r_y^{(2)}(q)\} \end{bmatrix}, \Psi_z \equiv \begin{bmatrix} \{r_x^{(3)}(q)\} \\ \{r_y^{(3)}(q)\} \end{bmatrix}$ where $q=(x,y)$ is the image position of the tracked point in the reference image and

the three point rotational flows $r^{(1)}(x,y), r^{(2)}(x,y), r^{(3)}(x,y)$ are defined by

$[r^{(1)}, r^{(1)}, r^{(1)}] \equiv \left[\begin{pmatrix} -xy \\ -(1+y^2) \end{pmatrix}, \begin{pmatrix} 1+x^2 \\ xy \end{pmatrix}, \begin{pmatrix} -y \\ x \end{pmatrix}\right]$ and where $\Psi_{\mathrm{I}x} \equiv -\{\nabla I \cdot \mathrm{r}^{(1)}(\mathrm{p})\}, \Psi_{\mathrm{I}y} \equiv -\{\nabla I \cdot \mathrm{r}^{(2)}(\mathrm{p})\}, \Psi_{\mathrm{I}z} \equiv -\{\nabla I \cdot \mathrm{r}^{(3)}(\mathrm{p})\}$, and where p gives the image coordinates of the pixel positions, and

where Λ is an $N_I \times 2N_L$ matrix where each row corresponds to a different image i and equals $\left[\{\mathrm{P}_U \cdot \delta\mathrm{A}^i\}^T \{\mathrm{P}_L \cdot \delta\mathrm{A}^i\}^T\right]$ where P_U and P_L are a unit 3-vector projecting onto two directions $\mathrm{A}_l \times (\hat{z} \times \mathrm{A}_l)$ and $\hat{z} \times \mathrm{A}_l$ which we refer to respectively as the upper and lower directions, where $\delta\mathrm{A}^i$ is the line flow $\delta\mathrm{A}_l^i \equiv \mathrm{A}_l^i - \mathrm{A}_l^0$, and $\hat{z}$ is the unit vector in the z direction and where S is a $N_i \times 2N_p$ matrix where each row corresponds to a different image i and equals $\left[\{s_x^i\}^T \{s_y^i\}^T\right]$, where $s_m^i \equiv \mathrm{q}_m^i - \mathrm{q}_m^0$ and denotes the image displacement for the m-th tracked point and where Δ is a $N_I \times N_X$ matrix, where each row corresponds to an image i and equals $\{\Delta I^i\}^T$ where ΔI^i is the change in image intensity with respect to the reference image and where I^i denotes the i-th intensity image and where $I_n^i = I^i(\mathrm{p}_n)$ denotes the image intensity at the n-th pixel position in I^i and where the notation $\{V\}$ is used to denote a vector with elements given by the V^a.

11. The method of claim 8, wherein step (d) comprises:

computing the best rank-3 factorization of $\overline{\mathrm{D}}_{CH} \approx M^{(3)} S^{(3)T}$ where $M^{(3)}, S^{(3)}$ are rank 3 matrices corresponding respectively to motion and structure, using SVD.

12. The method of claim 8, wherein step (e) comprises:

eliminating structure unknowns Q_z, B_z and Z^{-1} from the $\overline{\Phi}_a$ to get $3N_{tot}$ linear constraints on the U and Ω using the linear equation, $[\overline{\Phi}_x\ \overline{\Phi}_y\ \overline{\Phi}_z] = \mathrm{H}^T S^{(3)} U + [\overline{\Psi}_x\ \overline{\Psi}_y\ \overline{\Psi}_z]\Omega$, where U and Ω are unknown 3x3 matrices, and $\overline{\Phi}$ and $\overline{\Psi}$ represent total translational and rotational flow vectors respectively, and solving these constraints with $O(N_{tot})$ computations using the SVD, where the total translational flow vectors are defined by $\overline{\Phi}_x \equiv \begin{bmatrix} \Phi_x \\ \omega_{\mathrm{L}} \Phi_{\mathrm{L}x} \\ \omega_{\mathrm{I}} \Phi_{\mathrm{I}x} \end{bmatrix}$ and similarly for the y and z components, and

$\Phi_{Lx} \equiv \begin{bmatrix} \{A_x B_U\} \\ \{A_x B_L\} \end{bmatrix}, \Phi_{Ly} \equiv \begin{bmatrix} \{A_y B_U\} \\ \{A_y B_L\} \end{bmatrix}, \Phi_{Lz} \equiv \begin{bmatrix} \{A_z B_U\} \\ \{A_z B_L\} \end{bmatrix}$ where A is the normal to a first plane that passes through the center of projection and the imaged line in the reference image, and the

normal to a second plane, B is defined by requiring $B \cdot A = 0$ and $B \cdot Q = -1$ for any point Q on the 3D line L,

and where $B_U \equiv B \cdot P_U$ and $B_L \equiv B \cdot P_L$ are the upper and lower components of B, and where

$$\Phi_x \equiv -\begin{bmatrix} \{Q_z^{-1}\} \\ \{0\} \end{bmatrix}, \Phi_y \equiv -\begin{bmatrix} \{0\} \\ \{Q_z^{-1}\} \end{bmatrix}, \Phi_z \equiv -\begin{bmatrix} \{q_x Q_z^{-1}\} \\ \{q_y Q_z^{-1}\} \end{bmatrix}$$

and where

$\Phi_{Ix} \equiv -\{Z^{-1}I_x\}, \Phi_{Iy} \equiv -\{Z^{-1}I_y\}, \Phi_{Iz} \equiv \{Z^{-1}(\nabla I \cdot \mathrm{p})\}$, and

where Q is the 3d coordinate for a tracked 3D point in the reference image, and Z_n is the depth from the camera to the 3D point imaged at the n-th pixel along the cameras optical axis; and

recovering the structure unknowns Q_z, B_z and Z^{-1} from

$\left[\overline{\Phi}_x\ \overline{\Phi}_x\ \overline{\Phi}_x\right] = \mathrm{H}^T S^{(3)} U + \left[\overline{\Psi}_x\ \overline{\Psi}_y\ \overline{\Psi}_z\right]\Omega$, given U and Ω.

13. The method of claim 8, wherein step (f) comprises:

using $S^{(3)}U \approx \left[\overline{\Phi}_x\ \overline{\Phi}_y\ \overline{\Phi}_z\right]$ and

$\overline{\mathrm{D}}_{CH} \approx C^{-1/2}\{T_x\}\overline{\Phi}^T \mathrm{H}^{\mathrm{T}} + C^{-1/2}\{T_y\}\overline{\Phi}_y^T \mathrm{H}^{\mathrm{T}} + C^{-1/2}\{T_z\}\overline{\Phi}_z^T \mathrm{H}^{\mathrm{T}}$ to recover the translations, and

recovering the rotations $\omega_x^i, \omega_y^i, \omega_z^i$ from

$\omega_x^i \overline{\Psi}_{xn} + \omega_{yx}^i \overline{\Psi}_{yn} + \omega_{zx}^i \overline{\Psi}_{zn} = C^{-1/2}\overline{\mathrm{D}}_n - \left(C^{-1/2}\left(\{T_x\}\overline{\Phi}^T + \{T_y\}\overline{\Phi}_y^T + \{T_z\}\overline{\Phi}_z^T\right)\right)_n^i$, wherein C is a constant $(\mathrm{N_I}-1)$ x $(\mathrm{N_I}-1)$ matrix with $C_{ii'} \equiv \delta_{ii'} + 1$ and T represents the translation.

Add 93

Add 101